

## Positive Time-Frequency Distributions Based on Joint Marginal Constraints

Javier Rodríguez Fonollosa

**Abstract**—This correspondence studies the formulation of members of the Cohen–Posch class of positive time-frequency energy distributions. Minimization of cross-entropy measures with respect to different priors and the case of no prior or maximum entropy were considered. It is concluded that, in general, the information provided by the classical marginal constraints is very limited, and thus, the final distribution heavily depends on the prior distribution. To overcome this limitation, joint time and frequency marginals are derived based on a “direction invariance” criterion on the time-frequency plane that are directly related to the fractional Fourier transform.

### I. INTRODUCTION

Time-frequency representations (TFR’s) are functionals that attempt to express the joint distribution of energy density in the time-frequency plane. There are several basic requirements that TFR’s need to satisfy, although the most common representations cannot satisfy all of them simultaneously. The first, and most obvious, of the desired properties is the positivity. Only positive functionals can be properly interpreted as distributions in the probabilistic sense. Second, a true joint distribution of time and frequency must satisfy what are known as the “marginal” requirements, i.e., when the joint distribution is integrated along one of the variables, it should yield the marginal distribution of the other variable. In the case of time-frequency analysis, the marginal distributions in time and frequency are  $|s(t)|^2$  and  $|S(\omega)|^2$ , respectively. Denoting the time frequency distribution as  $Q(t, \omega)$ , these requirements are formulated as

$$\text{Positivity:} \quad Q(t, \omega) > 0 \quad (1)$$

$$\text{Marginal Requirements:} \quad \int Q(t, \omega) dt = |S(\omega)|^2; \quad (2)$$

$$\int Q(t, \omega) d\omega = |s(t)|^2$$

Historically, time-frequency “distributions” were defined using different approaches [1]. Among them, the Spectrogram, the Wigner–Ville or, in general, bilinear distributions included in the Cohen general class are the most commonly used. Representations with signal-independent kernels are bilinear with respect to the signal under analysis. This bilinearity of the TFR implies that the positive and marginal requirements cannot be satisfied simultaneously unless we restrict the signal under analysis to the generalized Gaussian class. Efforts toward the development of kernels according to different optimality criteria (other than positivity) resulted in the formulation of “signal-dependent” representations [2], [3].

Nevertheless, “proper” distributions, i.e., time-frequency distributions that satisfy (1) and (2), do exist and can be easily generated [4]. They belong to what has been called the Cohen–Posch class of positive

distributions, all of which can be parameterized by a function  $\rho(x, y)$ . There is an infinite number of such distributions, and the problem lies in finding the most useful for any particular application.

In this correspondence, we consider an information theoretic approach for the definition of proper time-frequency distributions. It follows the pioneering work of Loughlin *et al.* [5]–[7]. We will show that sometimes, the two classical marginal requirements used as constraints in an entropy optimization problem cannot provide sufficient information to yield a representation with improved resolution with respect to the prior. In addition, in this case, the resulting distribution is very sensitive with respect to the time-frequency “orientation” of the signal under analysis.

### II. BACKGROUND: MCE-TFD

Recently, a member of Cohen–Posch class of positive distributions was defined by Loughlin *et al.* [5]–[7]. It was formulated as the solution to a minimum cross-entropy (MCE) optimization problem subject to linear constraints in general and, in particular, the fulfillment of the marginal conditions. That is, given a prior estimate  $\hat{P}(t, \omega) > 0$  of the unknown distribution  $Q(t, \omega)$ , it is obtained as

$$Q(t, \omega) = \arg \min_{Q(t, \omega)} \left[ \int \int \hat{Q}(t, \omega) \log \frac{\hat{Q}(t, \omega)}{\hat{P}(t, \omega)} dt d\omega \right] \quad (3)$$

subject to the marginal requirements of (2).

The Spectrogram or a combination of Spectrograms was chosen as a *a priori* distribution  $\hat{P}(t, \omega)$ . As shown in the following examples, in some cases, the *a priori* distribution was very much preserved in the outcome, whereas in some other cases, the result was highly independent of the prior. High similarity with respect to the prior is not necessarily negative but, in our opinion, should not depend on the time-frequency orientation of the signal under analysis.

Let us first illustrate the main limitations associated with the MCE-TFD as defined in [5]–[7] when only the marginals are used as constraints. MCE subject to the two marginal constraints in general can only improve, to a limited extent, the *a priori* distributions, unless signal components are oriented in “vertical” or “horizontal” lines in the time-frequency plane. In this case, MCE minimization w.r.t. most meaningful priors would yield basically the same result since the outcome is primarily determined by the constraints. As an example, consider two Gaussian envelope complex exponentials of different constant frequency and a time-frequency “rotated” version of them. The Spectrogram<sup>1</sup> of these signals together with the MCE-TFD with Spectrogram prior are illustrated in Figs. 1 and 2: Fig. 1(a) and (b) for the constant frequencies case and Fig. 2(a) and (b) for the linear frequencies case. The high “orientation” dependency of the MCE-TFD is manifest since a much higher resolution is obtained in the first case compared with the second.

The limitations of the MCE-TFD using the marginals constraints only can be easily recognized if we analyze the classical form of the solution to the cross-entropy minimization problem. Using the method of Lagrange multipliers [7]

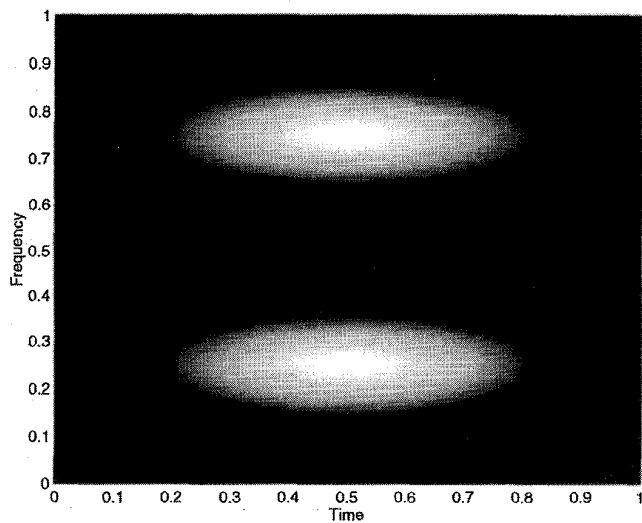
$$Q(t, \omega) = \hat{P}(t, \omega) \exp \{ -[1 + \alpha(\omega) + \beta(t)] \} \\ = \hat{P}(t, \omega) A(\omega) B(t) \quad (4)$$

<sup>1</sup>All distributions for synthetic signals have been computed using 256 samples of the signal and with a  $256 \times 256$  resolution. The Spectrogram employed a Gaussian window of variance equal to  $256/3.5$ .

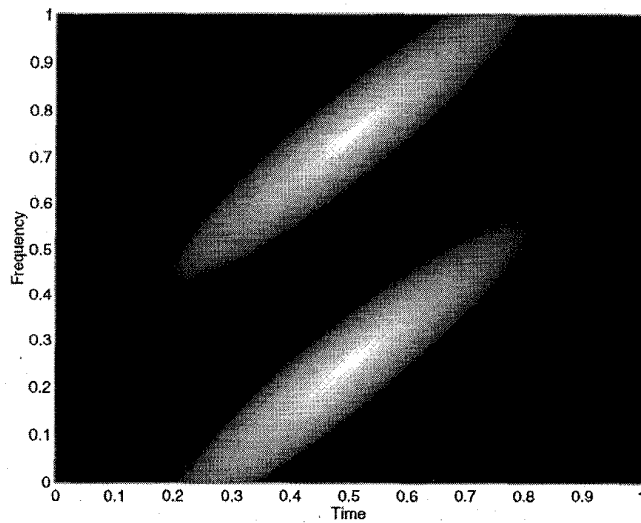
Manuscript received February 24, 1995; revised January 30, 1996. This work was supported by the National Plan of Spain, CICYT, grant TIC96-0500-C10-01, and the Generalitat de Catalunya, CIRIT, GRQ93-3021. Portions of this work were presented at ICASSP’94, Adelaide, South Australia, Australia, April 1994. The associate editor coordinating the review of this paper and approving it for publication was Dr. Patrick Flandrin.

The author is with the Universitat Politècnica de Catalunya, Department of Signal Theory and Communications, Campus Nord, D-5, C/Gran Capità s/n, 08034, Barcelona, Spain.

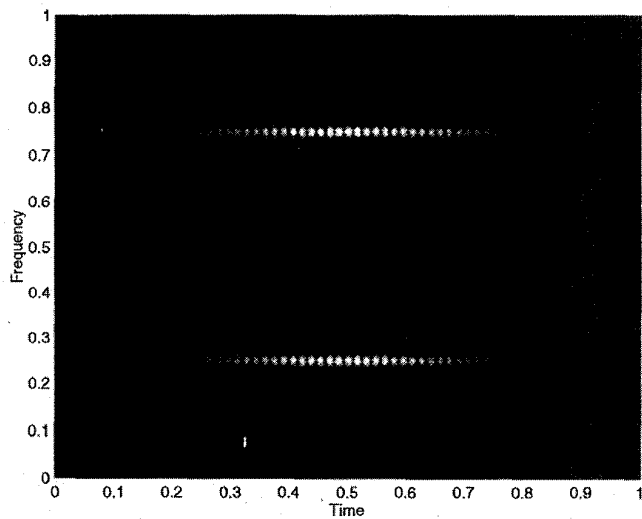
Publisher Item Identifier S 1053-587X(96)05272-5.



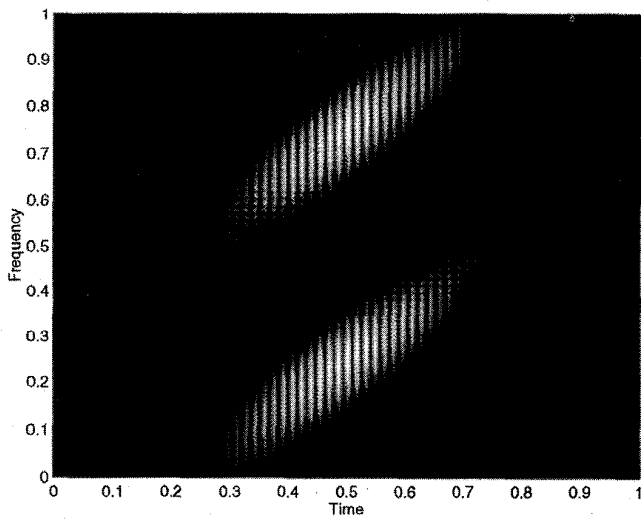
(a)



(a)



(b)



(b)

Fig. 1. (a) Spectrogram for two complex exponentials of constant frequency. (b) MCE-TFD for two complex exponentials of constant frequency.

Fig. 2. (a) Spectrogram for two linear chirp signals. (b) MCE-TFD for two linear chirp signals.

where  $\alpha$  and  $\beta$  are the Lagrange multipliers that need to be determined using the constraints. The resulting distribution is nothing but the *a priori* distribution shaped by a separable kernel  $A(\omega)B(t)$ .

As proposed by Loughlin *et al.*, a possible alternative to add more constraints, i.e., “information,” in the entropy optimization problem is incorporating the higher order moments of the time-frequency energy density such as the instantaneous frequency and bandwidth or group delay and time spread. The form of the solution then becomes

$$Q(t, \omega) = \hat{P}(t, \omega) \exp \{ -[1 + \alpha_0(\omega) + \beta_0(t) + \alpha_1(\omega)t + \beta_1(t)\omega + \alpha_2(\omega)t^2 + \beta_2(t)\omega^2] \}. \quad (5)$$

Much more information can be incorporated into the prior in this case since the exponential is no longer a separable function in the product of “time only” and “frequency only” functions. There are, however, two serious problems associated with these additional constraints: First, there is no general agreement about what the first and second moments of  $Q(t, \omega)$  should be, and second, numerical solution of the entropy optimization problem with the moments

constraints is not as simple as with the marginals constraints only [7].

We consider a different approach to supplement the “separable” information, provided we have the classical constraints. We propose the definition of joint time frequency marginals along arbitrary directions in the time-frequency plane to include them as constraints in the optimization problem. The purpose of this correspondence is to analyze the new “information” provided by the constraints in addition to the prior. Therefore, no analysis on the selection of the prior is included, but on the contrary, emphasis is given to building distributions based on constraints only.

### III. JOINT MARGINALS PROPERTIES

We investigated the possibility of defining joint generalized marginals along arbitrary directions of the time-frequency domain and concluded that the fractional or angular Fourier transform (AFT) (see [8]–[11] and references therein) could provide much insight into this problem. The AFT was defined independently by different

researchers, both mathematically and using physical considerations, and all definitions were later shown to be equivalent. The reader is referred to [11] for a thorough review of its definition, properties, and relation to the chirp and wavelet transform and applications. The AFT decomposes the signal under analysis using an orthonormal base of linear chirps and thus preserves the Parseval relation. The "orientation" of the basis functions (chirps) in the time-frequency plane depends on the angle  $\alpha$ . Particular cases of the AFT for  $\alpha = 0$  and  $\alpha = \pi/2$  are the identity operator and the Fourier transform, respectively. Precisely, the AFT at angle  $\alpha$  of a signal  $s(t)$  is [9]

$$\begin{aligned} S_\alpha(u) &= \int s(t) K_\alpha(t, u) dt \\ &= F_\alpha[s(t)] \end{aligned} \quad (6)$$

where

$$K_\alpha(t, u) = \begin{cases} \delta(t - u), & \text{if } \alpha \text{ is multiple of } 2\pi \\ \delta(t + u), & \text{if } \alpha + \pi \text{ is multiple of } 2\pi \\ \sqrt{\frac{1 - j \cot \alpha}{2\pi}} \exp\left(j \frac{t^2 + u^2}{2} \cot \alpha - jut \csc \alpha\right) & \text{if } \alpha \text{ is not a multiple of } \pi \end{cases} \quad (7)$$

where  $\csc$  and  $\cot$  are the *cosecant* and *cotangent* functions, respectively. In addition, it is additive with respect to the angle  $\alpha$ ,  $F_\beta\{F_\alpha[s(t)]\} = F_{\alpha+\beta}[s(t)]$ . Within a time-frequency analysis context, it is important to emphasize its relation to the Wigner-Ville Distribution (WD). In particular, defining the rotation operator  $R_\phi$  for 2-D functions, corresponding to a counterclockwise rotation by  $\phi$ , then

$$\text{WD}\{F_\alpha[s(t)]\} = R_\alpha\{\text{WD}[s(t)]\} \quad (8)$$

that is, the AFT at angle  $\alpha$  of the signal is equivalent to a  $\alpha$ -degree rotation of its WD.

Considering all the properties of the AFT and the fact that there is a general agreement on what the time-frequency representation of chirp signals should be, we concluded that the same integration of the distribution in the frequency and time axes should give the squared magnitude of the signal and its Fourier transform, as indicated by (2). The integration of the distribution along a rotated axis should give the squared magnitude of the AFT. That is

$$\int_t \int_\omega Q(t, \omega) \delta(t \cos \alpha + \omega \sin \alpha - u) dt d\omega = |S_\alpha(u)|^2. \quad (9)$$

It is well known that these generalized marginals are satisfied by the WD [9], [10]. In the next section, we will investigate the consequences of considering these joint marginals as constraints in an entropy-optimization problem.

#### IV. JOINT MARGINALS AS CONSTRAINTS

Apparently, a subclass of the Cohen-Posch class of positive distributions could be defined as the positive distributions that satisfy (9) for all  $\alpha$ ,  $-\pi/2 < \alpha \leq \pi/2$ . Nevertheless, (9) can be readily recognized as the Radon transform [12] of the distribution  $P(t, \omega)$ , and the inverse Radon transform theorem establishes that a 2-D function is uniquely determined from its Radon transform. Therefore,

the WD is the only distribution that satisfies (9) and, since the WD is a nonpositive functional in general, the subclass of the positive distributions satisfying (9) for all  $\alpha$ ,  $-\pi/2 < \alpha \leq \pi/2$  is empty.

Satisfaction of the joint marginals for all angles is incompatible with positivity, but this fact does not disqualify them as desirable properties. The same way the WD satisfies the joint marginals and is not positive everywhere, we can think of distributions that are positive and do not satisfy the joint marginals everywhere, but only at some angles. Hence, we propose to further restrict the Cohen-Posch class of positive distributions, enforcing the joint marginal constraints for some angles  $\alpha_i$  only. From an information theoretic point of view, the two classical marginal constraints of (2) supply information from the observation of the distribution at angles 0 and  $\pi/2$ . As seen in the case of MCE-TFD with Spectrogram prior, not much information is provided along other directions and the posterior, which is the resulting distribution, looks much like the prior when observed from those angles since it is modeled by a separable time-frequency kernel. On the other hand, satisfaction of the generalized marginals (9) at a limited number of angles can restrict the positive class of distributions and eventually make the outcome insensitive to the choice for the *a priori* distribution.

Therefore, we define a positive time-frequency distribution (ME-TFD) based on a maximum entropy criterion and given by

$$Q(t, \omega) = \arg \max_{\hat{Q}(t, \omega)} \left[ - \int \int \hat{Q}(t, \omega) \log \hat{Q}(t, \omega) dt d\omega \right] \quad (10.a)$$

subject to

$$\int_t \int_\omega Q(t, \omega) \delta(t \cos \alpha_i + \omega \sin \alpha_i - u) dt d\omega = |S_{\alpha_i}(u)|^2 \quad (10.b)$$

for  $\alpha_1 = 0$ ,  $\alpha_2 = \pi/2$ , and some additional  $\alpha_i$  to a total of  $N$  constraints. The information introduced by the joint marginals can be observed in the form of the solution to the entropy maximization problem (which is equivalent to cross-entropy minimization with uniform prior). Again, using the method of Lagrange multipliers (see the Appendix)

$$Q(t, \omega) = \exp \left\{ - \left[ 1 + \sum_{i=1}^N \lambda_i(t \cos \alpha_i + \omega \sin \alpha_i) \right] \right\} \quad (11)$$

where  $\lambda_i(t \cos \alpha_i + \omega \sin \alpha_i)$  are the Lagrange multipliers that result from the enforcement of the constraints. Note that  $\alpha_1 = 0$ , and  $\alpha_2 = \pi/2$ , and thus,  $\lambda_1$  and  $\lambda_2$  are a function of  $t$  and  $\omega$  only, respectively, whereas the others depend of a combination of both variables. It is apparent that the information introduced by the constraints is given (in general) in the form of a nonseparable kernel. Of course, if a reliable estimate of the distribution was available, it could be introduced as a prior, and the solution to the minimum cross-entropy problem would have the form

$$Q(t, \omega) = \hat{P}(t, \omega) \exp \left\{ - \left[ 1 + \sum_{i=1}^N \lambda_i(t \cos \alpha_i + \omega \sin \alpha_i) \right] \right\}. \quad (12)$$

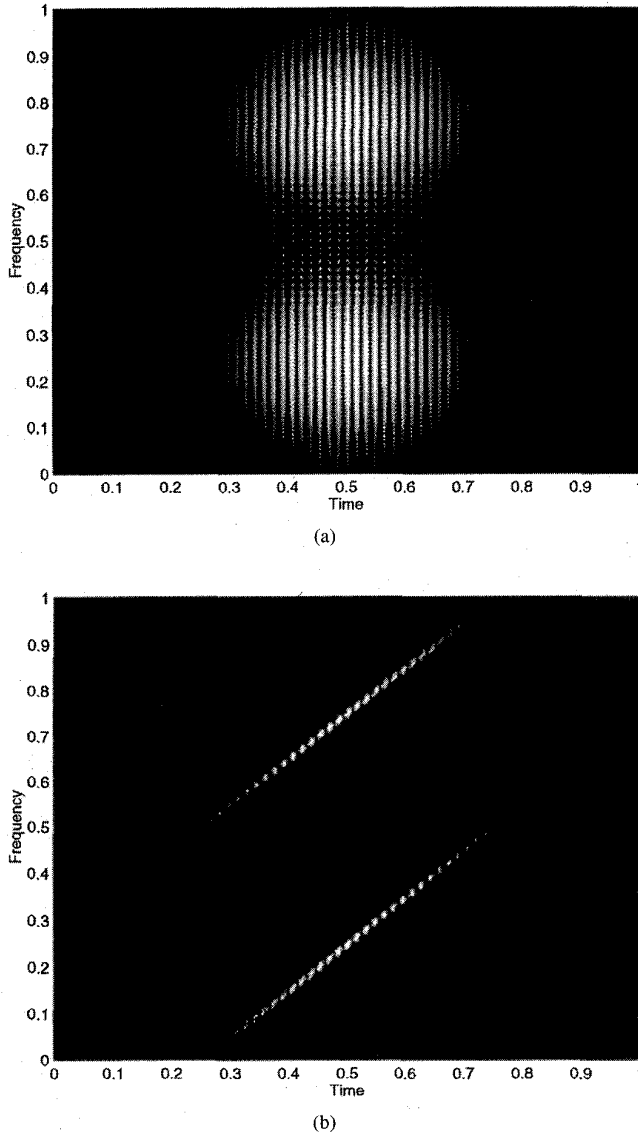


Fig. 3. (a) ME-TFD with the marginal constraints only. (b) ME-TFD with two additional constraints.

No proof of consistency can be given for satisfaction of the marginals for a given set of angles. From a theoretical point of view, the satisfaction of the joint marginals should be regarded as a desirable property that might be incompatible with positivity. From a practical point of view, no inconsistency has been observed using a small number of constraints. Analysis of the robustness of the method with respect to slight rotations in the angles is deferred for a later publication.

## V. IMPLEMENTATION

Computation of (10) is a multivariate constrained maximization problem that can be solved using the Lagrange multipliers method but for which a closed-form solution cannot be obtained in general. Yet, there is an iterative solution to the problem in which each of the univariate optimization problems is solved independently [14], [15]. The same technique was used in [7] for the MCE-TFD. In our case,

we start with the uniform distribution (instead of the Spectrogram)

$$Q_0^0(t, \omega) = \frac{1}{E_s};$$

$$E_s = \int |s(t)|^2 dt \quad (13)$$

and follow with the  $N$  individual constraint enforcement iterations

$$Q_i^0(t, \omega) = \frac{Q_{i-1}^0(t, \omega) |S_{\alpha_i}(u_i)|^2}{\int_{t'} \int_{\omega'} Q_{i-1}^0(t', \omega') \delta[t' \cos(\alpha_i) + \omega' \sin(\alpha_i) - u_i] dt' d\omega'}$$

for  $i = 1, \dots, N$  and  $u_i = t \cos(\alpha_i) + \omega \sin(\alpha_i)$  (14)

for  $\alpha_1 = 0$ ,  $\alpha_2 = \pi/2$ , and the desired additional  $\alpha_i$ . Computation of the AFT is not directly required since  $|S_{\alpha}(u)|^2$  can be obtained as the integration along the corresponding line of the WD. Please note that since we are forcing each of the constraints independently, after each iteration, only the last one (and perhaps none of the other constraints) will be satisfied. Then, using  $Q_N^0(t, \omega)$  as prior, we obtain  $Q_1^1(t, \omega), \dots, Q_N^1(t, \omega)$ . This procedure should be repeated until convergence is attained to the desired precision

$$Q(t, \omega) = \lim_{k \rightarrow \infty} Q_N^k(t, \omega). \quad (15)$$

Convergence is, however, assured, provided the constraints are consistent, by the convexity of the cross-entropy cost function [13], which implies

$$H[Q(t, \omega), Q_{i+1}^k(t, \omega)] \leq H[Q(t, \omega), Q_i^k(t, \omega)] \quad (16)$$

where  $H[Q(t, \omega), P(t, \omega)]$  is the cross-entropy between  $Q(t, \omega)$  and  $P(t, \omega)$ .

Addition of constraints only increases the computational complexity of the method linearly.

## VI. EXAMPLES

In our simulations, we considered two additional values of  $\alpha_i$  in addition to the "horizontal" and "vertical" marginals  $\alpha_3 = \pi/4$  and  $\alpha_4 = -\pi/4$ . These marginals are very simple to compute since, in a discrete implementation, they correspond to the sum of the diagonals of the matrix that represents the WD. This choice of angles  $\alpha_i$  means that the form of the distribution, given by the particularization of (11), is

$$Q(t, \omega) = \exp \{ -[1 + \lambda_1(t) + \lambda_2(\omega) + \lambda_3(t + \omega) + \lambda_4(t - \omega)] \}. \quad (17)$$

### A. Synthetic Signals

We first illustrate how the results obtained using the ME-TFD are consistent with what should be expected from its theory. Fig. 3(a) and (b) show the distributions obtained using ME-TFD with the two usual marginal constraints only and with the addition of the joint constraints at angles  $\alpha = \pi/4$  and  $\alpha = -\pi/4$ . The "orientation" dependency of the MCE-TFD has been diminished with the addition of only two joint constraints as indicated by (17) compared with (4). In addition, as expected, a resolution similar to the one obtained with the MCE-TFD in the constant frequency case (Fig. 1(b)) is obtained using the ME-TFD with the linear chirp.

### B. Real Signals

We illustrate the application of the distributions described in this correspondence to a natural signal. The signal under analysis is a chirp-like signal recorded from a dolphin. In Fig. 4(a), we give the

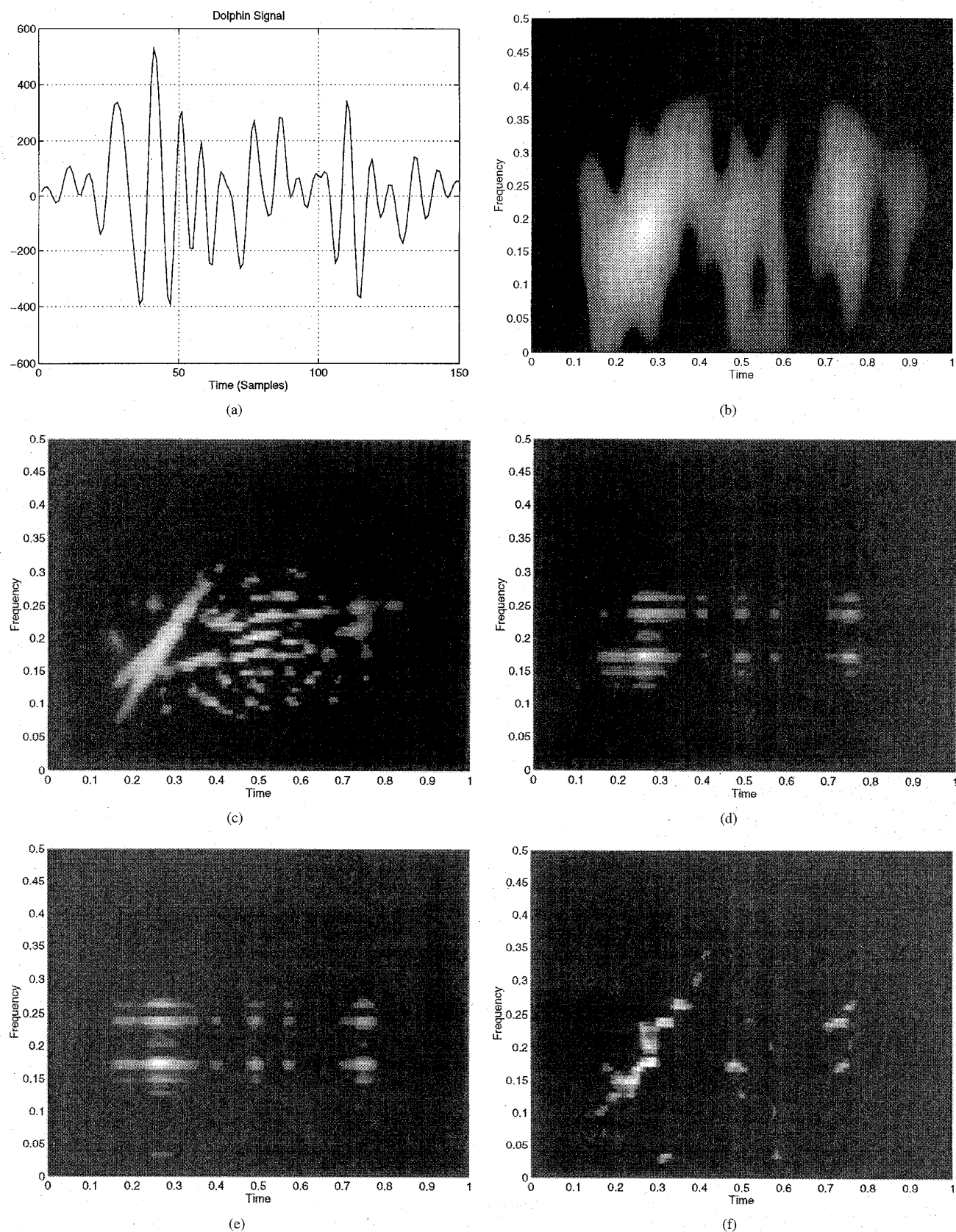


Fig. 4. (a) Temporal representation of the dolphin signal. (b) Spectrogram of the dolphin signal. (c) Wigner-Ville distribution of the dolphin signal. (d) MCE-TFD of the dolphin signal with the two marginals as constraints. (e) ME-TFD of the dolphin signal with the two marginals as constraints. (f) ME-TFD of the dolphin signal with four marginals as constraints.

temporal representation of this signal. Fig. 4(b) and (c) show the Spectrogram and the WD of the dolphin signal, respectively. The Spectrogram is characterized by a poor time-frequency resolution, whereas the WD suffers from interference terms. In Fig. 4(d), the MCE-TFD employing the two classical marginal constraints only and Spectrogram prior is illustrated. If no prior is employed, i.e., if a maximum entropy TFD is computed using only the two marginals, the resulting distribution is very similar and is illustrated in Fig. 4(e). We can observe how the marginal constraints "shape" the *a priori* distribution along the horizontal and vertical directions only. Finally, the ME-TFD that makes use of the two classical marginals and two joint marginals is given in Fig. 4(f). A chirp-like representation is obtained without the interfering terms of the WD. It is fair to mention that the MCE-TFD could be potentially improved after a careful selection of the prior, as suggested in [7], but emphasis in this paper is given to the information supplied by the constraints and not to the selection of the prior.

## VII. CONCLUSIONS

A new set of joint constraints has been proposed for the class of positive time-frequency distributions. They are the equivalent, along nonhorizontal or vertical directions, of the two classical marginals and can be directly related to the fractional Fourier transform. Although the only distribution that satisfies the complete set of joint constraints is the Wigner-Ville Distribution, they can be partially used to restrain the Cohen-Posch class of positive distributions. They have been employed as additional constraints in the formulation of positive distributions solution to entropy optimization problems. The main advantage of the joint constraints when employed in entropy optimization problems is that the *a priori* distribution can be modified with a kernel that is not separable in time and frequency. Only time and frequency separable functions can modify the prior when using the two marginals as unique constraints in the same problem. Positive distributions defined as solution to a maximum entropy optimization subject to the classical plus two additional constraints have been analyzed and compared with existing approaches with both synthetic and real signals.

## APPENDIX

Given a prior distribution,  $\hat{P}(t, \omega)$  and a set of equality constraints

$$\int_t \int_\omega Q(t, \omega) \delta(t \cos \alpha_i + \omega \sin \alpha_i - u) dt d\omega = |S_{\alpha_i}(u)|^2 \quad (\text{A.1})$$

for  $\alpha_1 = 0$ ,  $\alpha_2 = \pi/2$ , and "some" additional  $\alpha_i$  to a total of  $N$  constraints, we wish to find the distribution that minimizes

$$Q(t, \omega) = \arg \min_{\hat{Q}(t, \omega)} \left[ \int \int \hat{Q}(t, \omega) \log \frac{\hat{Q}(t, \omega)}{\hat{P}(t, \omega)} dt d\omega \right]. \quad (\text{A.2})$$

Introducing the Lagrange multipliers  $\lambda_i(u_i)$  forming the expression [13]

$$\begin{aligned} & \int \int Q(t', \omega') \log \frac{Q(t', \omega')}{\hat{P}(t', \omega')} dt' d\omega' \\ & + \sum_{i=1}^N \int \lambda_i(u_i) \int \int Q(t', \omega') \\ & \cdot \delta(t' \cos \alpha_i + \omega' \sin \alpha_i - u_i) dt' d\omega' du_i. \end{aligned} \quad (\text{A.3})$$

Equating the variation of (A.3) with respect to  $Q(t, \omega)$  to zero

$$\log \frac{Q(t, \omega)}{\hat{P}(t, \omega)} + 1 + \sum_{i=1}^N \lambda_i(t \cos \alpha_i + \omega \sin \alpha_i) = 0 \quad (\text{A.4})$$

which, solving for  $Q(t, \omega)$ , leads to

$$Q(t, \omega) = \hat{P}(t, \omega) \exp \left\{ - \left[ 1 + \sum_{i=1}^N \lambda_i(t \cos \alpha_i + \omega \sin \alpha_i) \right] \right\}. \quad (\text{A.5})$$

This proves (12) while considering uniform prior (11). In addition, taking  $\alpha_1 = 0$ ,  $\alpha_2 = \pi/2$ ,  $\alpha_3 = \pi/4$ , and  $\alpha_4 = -\pi/4$  and uniform prior proves (17).

## ACKNOWLEDGMENT

The author expresses his gratitude to Prof. L. Cohen and the anonymous reviewers for their enlightening discussions and especially to reviewer A for her/his thorough review of the manuscript.

## REFERENCES

- [1] L. Cohen, "Time-frequency distributions—A review," *Proc. IEEE*, vol. 77, no. 7, pp. 941–981, 1989.
- [2] R. Baraniuk and D. Jones, "A signal-dependent time-frequency—I Representation: Optimal kernel design," *IEEE Trans. Signal Processing*, vol. 41, no. 4, pp. 263–284, 1993.
- [3] D. Jones and T. Parks, "A high resolution data-adaptive time-frequency representation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, no. 12, pp. 2127–2135, 1985.
- [4] L. Cohen and T. E. Posch, "Positive time-frequency distribution functions," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, no. 1, pp. 31–38, 1985.
- [5] P. J. Loughlin, J. W. Pitton, and L. E. Atlas, "An information-theoretic approach to positive time-frequency distributions," in *Proc. ICASSP-92*, San Francisco, CA, 1992, pp. V 125–128.
- [6] J. W. Pitton, P. J. Loughlin, and L. E. Atlas, "Positive time-frequency distributions via maximum entropy deconvolution of the evolutionary spectrum," in *Proc. ICASSP-93*, Minneapolis, MN, 1993, pp. IV 436–439.
- [7] P. J. Loughlin, J. W. Pitton, and L. E. Atlas, "Construction of positive time-frequency distributions," *IEEE Trans. Signal Processing*, vol. 42, no. 10, pp. 2697–2705, 1994.
- [8] B. W. Dickinson and K. Steiglitz, "Eigenvectors and functions of the discrete Fourier transform," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-30, no. 1, pp. 25–31, 1982.
- [9] L. Almeida, "An introduction to the angular Fourier transform," in *Proc. ICASSP-93*, Minneapolis, MN, 1993, pp. III 257–260.
- [10] R. G. Dorsch and A. W. Lohmann, "The fractional Fourier transform," in *Proc. ATHOS Workshop Syst. Identification High Order Stat.*, Sophia Antipolis, France, 1993, pp. 1–8.
- [11] H. M. Ozaktas and B. Barshan, "Convolution, filtering, and multiplexing in fractional Fourier domains and their relation to chirp and wavelet transforms," *J. Opt. Soc. Amer. A, Optics Image Sci. Vision*, vol. 11, no. 2, pp. 547–559, 1994.
- [12] S. R. Deans, *The Radon Transform and Some of Its Applications*. New York: Wiley, 1983.
- [13] J. E. Shore and R. W. Johnson, "Properties of cross-entropy minimization," *IEEE Trans. Inform. Theory*, vol. IT-27, no. 4, pp. 472–482, 1981.
- [14] Y. A. Alsaka, N. S. Tzannes, and W. A. Marinelli, "An efficient algorithm for implementing the relative entropy method," in *Proc. ICASSP-88*, 1988, pp. 2384–2387.
- [15] R. Burr, "Iterative convex I-projection algorithm for maximum entropy and minimum cross-entropy computations," *IEEE Trans. Inform. Theory*, vol. 35, no. 3, pp. 695–698, 1989.